



Open Science at the Princess Máxima Center

Judith Boer

Lygature Partnerships Meetup

29 October 2019

Rationale for Open Science

- Responsibility to make scientific data FAIR
findable, accessible, interoperable and re-usable
 - in our own research group
 - in collaborations with other researchers
 - in public repositories
- Scientific data should be easy to understand, allow follow-up studies and be shared to answer new questions

Motivating example: brain tumor data

Subject: genes to distinguish medulloblastoma and ependymoma

Date: September 4, 2015 at 2:11:05 PM PDT

To: m.l.denboer@erasmusmc.nl

Dear Dr. den Boer,

I would like to distinguish medulloblastoma from ependymoma using differentially expressed mRNA. In the attached paper you identified such mRNAs, the 394 and 582 probesets indicated in Figure 1.

It appears that the complete Affymetrix data and analysis was not part of the 2007 paper. I wonder if you could send me those data or point me to a public repository that contains the data.

Thanks for your help.

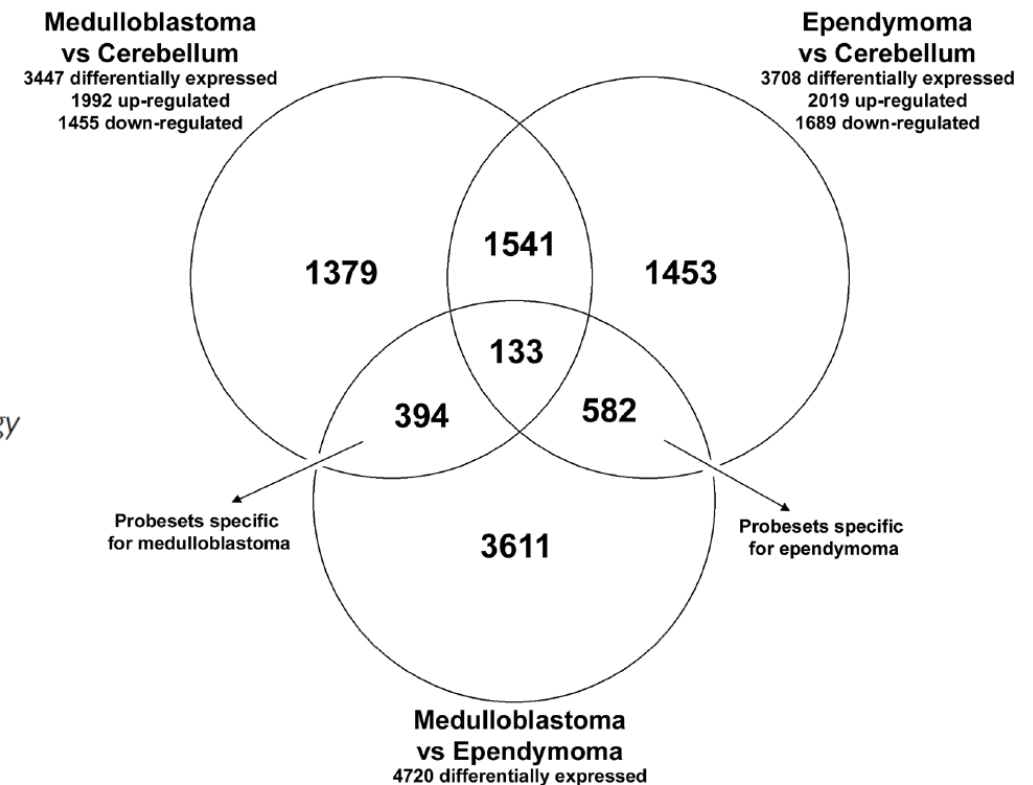
Motivating example: brain tumor data

Differential expression and prognostic significance of *SOX* genes in pediatric medulloblastoma and ependymoma identified by microarray analysis

Judith M. de Bont, Johan M. Kros, Monique M.C.J. Passier, Roel E. Reddingius, Peter A.E. Sillevs Smitt, Theo M. Luider, Monique L. den Boer, and Rob Pieters

Department of Pediatric Oncology and Hematology, Erasmus MC Sophia Children's Hospital, University Medical Center Rotterdam, Rotterdam (J.M.B., M.M.C.J.P., R.E.R., M.L.B., R.P.); Department of Pathology (J.M.K.) and Department of Neurooncology (P.A.E.S.S., T.M.L.), Erasmus MC, University Medical Center Rotterdam, Rotterdam; The Netherlands

Neuro Oncol. 2008



Motivating example: brain tumor data

The situation

- Samples were collected & microarrays run in 2005
- No supplemental data with lists of differentially expressed genes were submitted
- Arrays were not deposited in a public database
- Backups from the PhD student and technician were unstructured and incomplete

Motivating example: brain tumor data

The solution (>> 1 day!!)

- The PhD student sent the lists of differentially expressed genes, normalized data, and sample annotation
- Backup CDs contained the raw data files, but 4 files were corrupt
- Affymetrix cracked an old archive type “DTT” and recovered the corrupt raw data files

Motivating example: brain tumor data

FAIR at last

- Extra checks to make sure sample annotation and expression data were okay
- Sent the collaborating researcher the two lists of differentially expressed genes (Venn diagram Figure 1)
- Uploaded the raw data files, the original normalized data matrix and the meta-data to a public repository



Erasmus MC

2010

paper lab notebook

publication archiving

2015

submissions & publications including raw and processed data

experiments & raw data

electronic lab notebook

group data steward

2018



Princess Maxima Center

2019

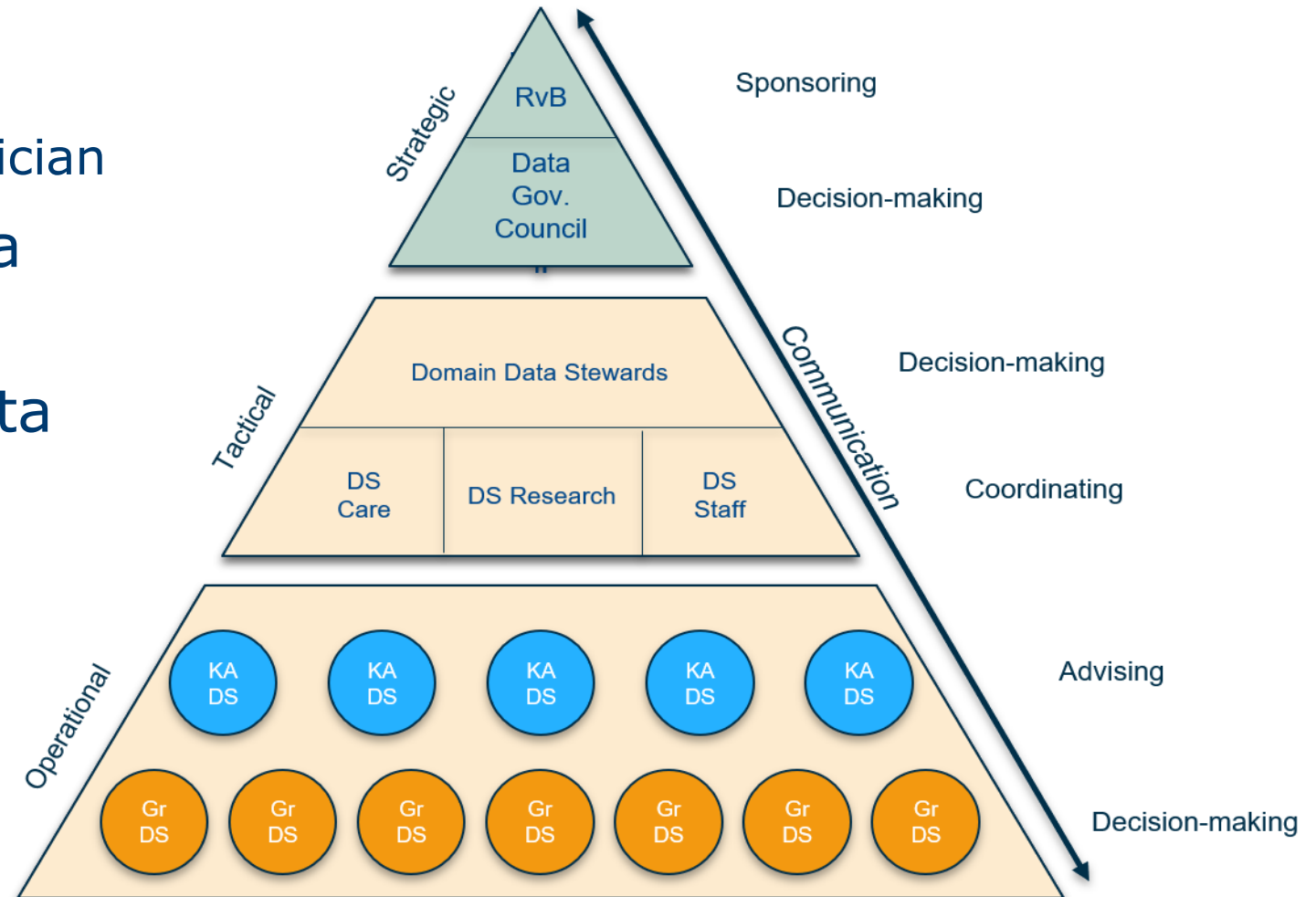


Data stewardship roles and responsibilities

- Group Data Steward in each research group
 - postdoc, technician, bioinformatician
- Understands the kind of data processed in his/her group
- Takes care of the group's data on an operational level

Abbreviations

Data Governance Council
Domain Data Stewards (for Research: PIs)
DS: Data Steward
KA: Knowledge Area
Gr: Group



Who is responsible for the data generated in a research group?

A The researcher who produced the data

B The group data steward

C The principle investigator / head of the group

Who is responsible for the data generated in a research group?

A The researcher who produced the data

→ experimental design, data quality, objective interpretation

B The group data steward

C The principle investigator / head of the group

Who is responsible for the data generated in a research group?

A The researcher who produced the data

B The group data steward

→ curated, findable, reusable group data

C The principle investigator / head of the group

Who is responsible for the data generated in a research group?

A The researcher who produced the data

B The group data steward

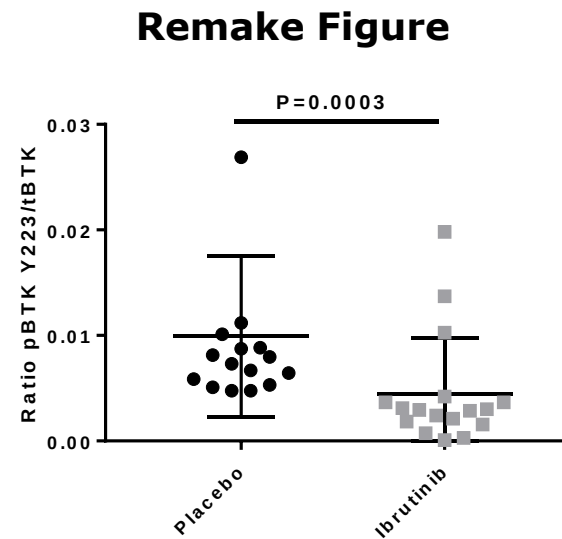
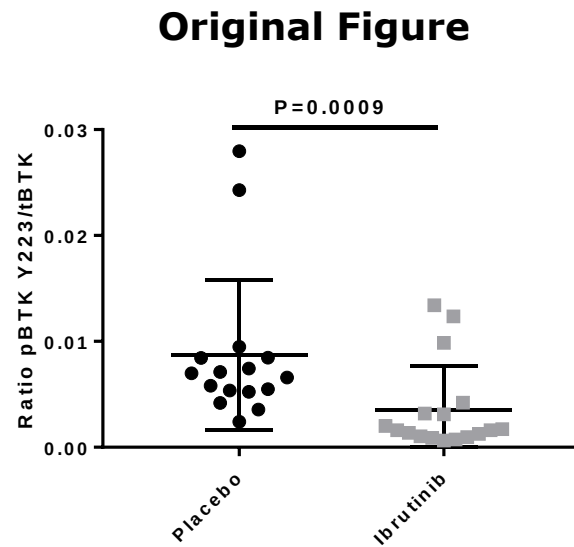
C The principle investigator / head of the group
→ takes tactic data-related decisions, e.g. reusability

Internal Open Science

- Raw and processed data, data formats, protocols
- (Electronic) Lab Notebook with links to experimental data
- Archiving for publications using standard format
- Finding and re-using data: including internal audits

Internal Open Science

- Raw and processed data, data formats, protocols
- (Electronic) Lab Notebook with links to experimental data
- Archiving for publications using standard format
- Finding and re-using data: including internal audits



- Make data available to collaborators & in public repositories
- Also publish negative results

High *PDGFRA* Expression Does Not Serve As An Effective Therapeutic Target In *ERG*-deleted B-cell Precursor Acute Lymphoblastic Leukemia

Isabel S. Jerchel, Danai Chatzivasileiou, Alex Q. Hoogkamer, Judith M. Boer, H. Berna Beverloo, Rob Pieters, Monique L. den Boer

Haematologica February 2018 103: e73-e77; **Doi:**10.3324/haematol.2017.171702

Altered expression of miR-24, miR-126 and miR-365 does not affect viability of childhood TCF3-rearranged leukemia cells

F Akbari Moqadam, J M Boer, E A M Lange-Turenhout, R Pieters & M L den Boer 

Leukemia **28**, 1008–1014 (2014) | [Download Citation](#) 

High STAP1 expression in DUX4-rearranged cases is not suitable as therapeutic target in pediatric B-cell precursor acute lymphoblastic leukemia

Elisabeth M. P. Steeghs, Marjolein Bakker, Alex Q. Hoogkamer, Judith M. Boer, Quirine J. Hartman, Femke Stalpers, Gabriele Escherich, Valerie de Haas, Hester A. de Groot-Kruseman, Rob Pieters & Monique L. den Boer 

Scientific Reports **8**, Article number: 693 (2018) | [Download Citation](#) 

Internal: example lab notebook

- Raw and processed data, data formats, protocols
- (Electronic) Lab Notebook with links to experimental data
- Archiving for publications using standard format
- Finding and re-using data: including internal audits

If you only implement 1 thing: Experiment number

- Electronic Lab Notebook defines structure and access
- Experiment number links experiment to stored data: **Date + Initials + Nr**


Projects


Studies


Experiments

Journal Inventory Procedures Supplies Configuration File Storage				
Dashboard Experiment Browser Timeline Projects Studies Experiment List				
Experiments Administration				
Filters: Femke Hormann b_Ch21 NUTM1 -- All Statuses -- -- Any Due Date --				
Experiment	Project ID	Study	Status	
20181217_FMH-115_RT-PCR_NUTM1 primers	b_Ch21	NUTM1	Completed	
20181210_FMH-111_ordering NUTM1 fusion primers	b_Ch21	NUTM1	Completed	
20181217_FMH-114_NUTM1 stain cytopspins	b_Ch21	NUTM1	Completed	
20181212_FMH-112_LaminB staining cytopspins	b_Ch21	NUTM1	Configuring	
20181206_FMH-110_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring	
20181203_FMH-108_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring	
20181126_FMH-105_RT-PCR_NUTM1 primers on products FMH-103	b_Ch21	NUTM1	Completed	

elabjournal

Chr21

Cordblood

Culture

Knockdown and knockout

NUTM1

STAT5 activation

Virus production

Filename

20181107_FMH-98_Ordering primers RT-PCR fusion detection

20181210_FMH-111_ordering NUTM1 fusion primers

20181217_FMH-114_NUTM1 stain cytopspins

20181217_FMH-115_RT-PCR_NUTM1 primers

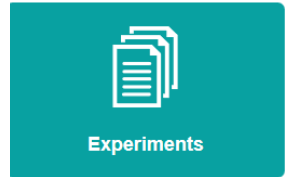
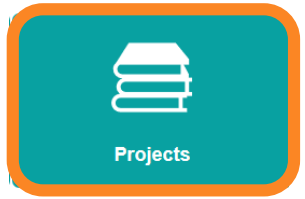
20181116_FMH-103_RT-PCR_NUTM1 primers

20181113_FMH-100_RT-PCR_NUTM1primers

20181212_FMH-112_LaminB staining cytopspins

If you only implement 1 thing: Experiment number

- Electronic Lab Notebook defines structure and access
- Experiment number links experiment to stored data



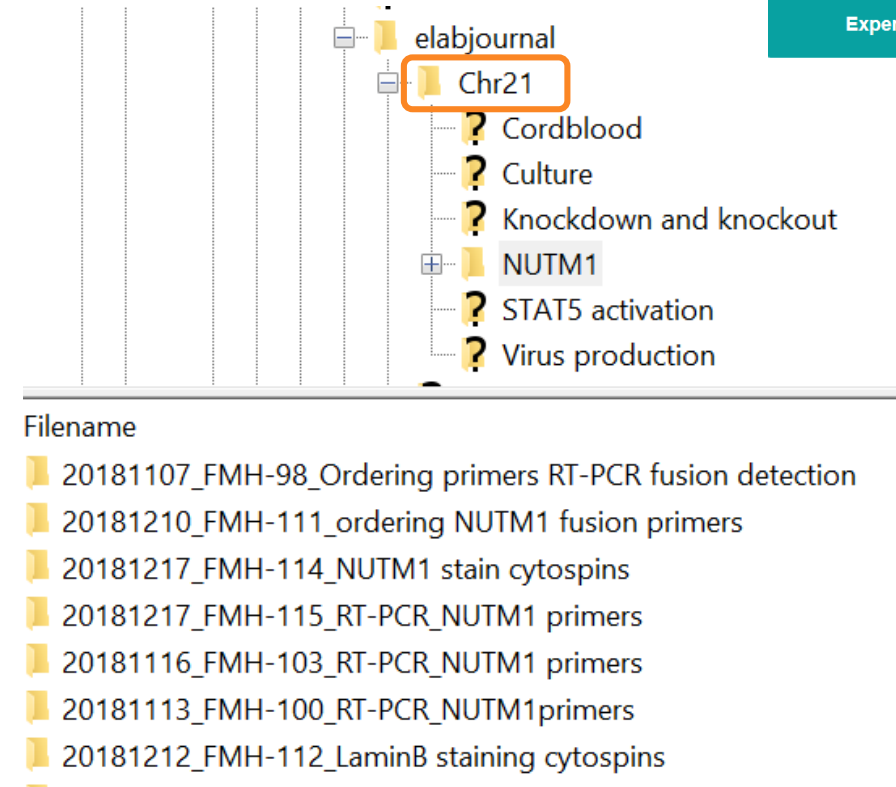
JournalInventoryProceduresSuppliesConfigurationFile Storage

DashboardExperiment BrowserTimelineProjectsStudiesExperiment List

Experiments Administration

Filters: Femke Hormann **b_Ch21** NUTM1 -- All Statuses -- -- Any Due Date --

Experiment	Project ID	Study	Status
20181217_FMH-115_RT-PCR_NUTM1 primers	b_Ch21	NUTM1	Completed
20181210_FMH-111_ordering NUTM1 fusion primers	b_Ch21	NUTM1	Completed
20181217_FMH-114_NUTM1 stain cytopspins	b_Ch21	NUTM1	Completed
20181212_FMH-112_LaminB staining cytopspins	b_Ch21	NUTM1	Configuring
20181206_FMH-110_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181203_FMH-108_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181126_FMH-105_RT-PCR_NUTM1 primers on products FMH-103	b_Ch21	NUTM1	Completed



If you only implement 1 thing: Experiment number

- Electronic Lab Notebook defines structure and access
- Experiment number links experiment to stored data



Projects



Studies



Experiments

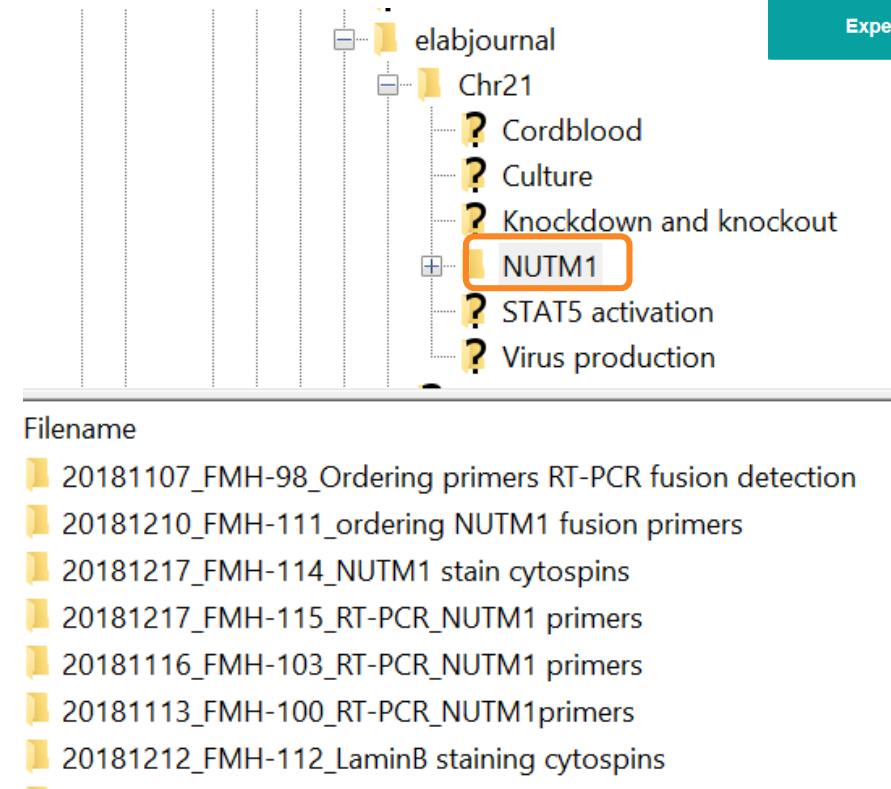
JournalInventoryProceduresSuppliesConfigurationFile Storage

DashboardExperiment BrowserTimelineProjectsStudiesExperiment List

Experiments Administration

Filters: Femke Hormann b_Ch21 **NUTM1** -- All Statuses -- -- Any Due Date --

Experiment	Project ID	Study	Status
20181217_FMH-115_RT-PCR_NUTM1 primers	b_Ch21	NUTM1	Completed
20181210_FMH-111_ordering NUTM1 fusion primers	b_Ch21	NUTM1	Completed
20181217_FMH-114_NUTM1 stain cytopspins	b_Ch21	NUTM1	Completed
20181212_FMH-112_LaminB staining cytopspins	b_Ch21	NUTM1	Configuring
20181206_FMH-110_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181203_FMH-108_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181126_FMH-105_RT-PCR_NUTM1 primers on products FMH-103	b_Ch21	NUTM1	Completed



If you only implement 1 thing: Experiment number

- Electronic Lab Notebook defines structure and access
- Experiment number links experiment to stored data



Projects



Studies



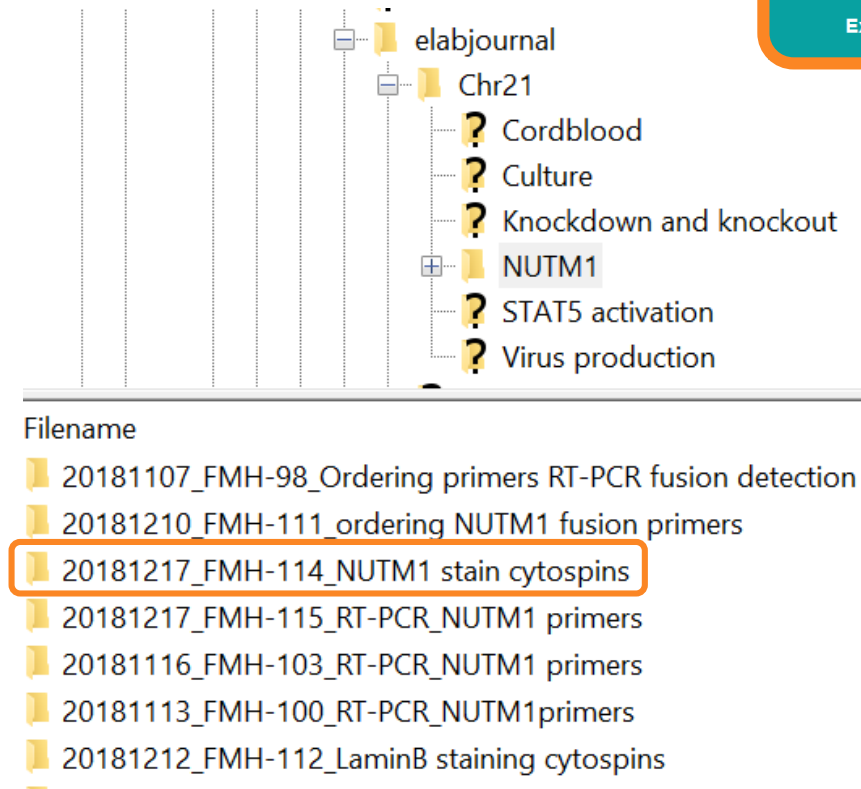
Experiments

Journal	Inventory	Procedures	Supplies	Configuration	File Storage
Dashboard	Experiment Browser	Timeline	Projects	Studies	Experiment List

Experiments Administration

Filters: Femke Hormann b_Ch21 NUTM1 -- All Statuses -- -- Any Due Date --

Experiment	Project ID	Study	Status
20181217_FMH-115_RT-PCR_NUTM1 primers	b_Ch21	NUTM1	Completed
20181210_FMH-111_ordering NUTM1 fusion primers	b_Ch21	NUTM1	Completed
20181217_FMH-114_NUTM1 stain cytopspins	b_Ch21	NUTM1	Completed
20181212_FMH-112_LaminB staining cytopspins	b_Ch21	NUTM1	Configuring
20181206_FMH-110_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181203_FMH-108_NUTM1 stain cytopspins	b_Ch21	NUTM1	Configuring
20181126_FMH-105_RT-PCR_NUTM1 primers on products FMH-103	b_Ch21	NUTM1	Completed



If you only implement 1 thing: Experiment number

- Electronic Lab Notebook defines structure and access
- Experiment number links experiment to stored data


Projects


Studies


Experiments

Journal

Inventory

Procedures

Supplies

Configuration

File Storage

Dashboard

Experiment Browser

Timeline

Projects

Studies

Experiment List

Experiments Administration

Filters: Femke Hormann b_Chr21 NUTM1 -- All Statuses -- -- Any Due Date --

Experiment	Project ID	Study	Status
20181217_FMH-115_RT-PCR_NUTM1 primers	b_Chr21	NUTM1	Completed
20181210_FMH-111_ordering NUTM1 fusion primers	b_Chr21	NUTM1	Completed
20181217_FMH-114_NUTM1 stain cytopspins	b_Chr21	NUTM1	Completed
20181212_FMH-112_LaminB staining cytopspins	b_Chr21	NUTM1	Configuring
20181206_FMH-110_NUTM1 stain cytopspins	b_Chr21	NUTM1	Configuring
20181203_FMH-108_NUTM1 stain cytopspins	b_Chr21	NUTM1	Configuring
20181126_FMH-105_RT-PCR_NUTM1 primers on products FMH-103	b_Chr21	NUTM1	Completed

elabjournal

Chr21

Cordblood

Culture

Knockdown and knockout

NUTM1

STAT5 activation

Filename

4823B_NUTM1002_metadata.pdf

8449_NUTM1005_metadata.pdf

20181205_FMH-108_isotype control_001_DAPI adjust.png

20181205_FMH-108_isotype control_001.png

20181217_FMH-114_NUTM1 stain cytopspins.pdf

20181218_NUTM1_Femke.tif

20181218_NUTM1_Femke.tif - 4823B_NUTM1002.tif

20181218_NUTM1_Femke.tif - 8449_NUTM1005.tif

Pictures.png

Pictures.pptx

Sample 1_001.png

Sample 1_002.png

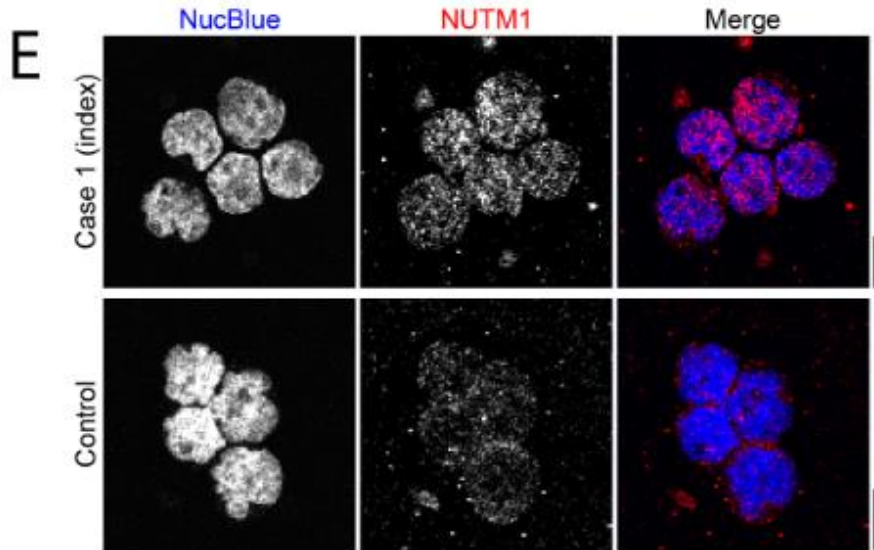
Sample 2_001.png

Sample 2_002.png

Sample 3_001.png

If you implement a 2nd thing: Archiving publications

- Little extra work if stored data is structured
- Read-me file for each figure referring to raw data and describing processing steps



E: Representative confocal microscope images of cytopins of the index case #1 with a SLC12A6-NUTM1 fusion and a B-other-ALL control sample immunostained with NUT antibody (NUTM1; red in merge).

- Data: Pictures from experiment 20181217_FMH-114_NUTM1 stain cytopins Used:
 - 20181218_NUTM1_Femke.lif - 4823B_NUTM1002-1.tif
 - 20181218_NUTM1_Femke.lif - 8449_NUTM1005-1.tif
- The intensity of the red channel was set to 1250. A representative cutout containing 4-5 cells was made. A color composite containing DAPI and the red channel was made. A grayscale figure was made for both the red and the blue channel.
- A scale bar was made in illustrator.
- This resulted in Imaging NUTM1.ai (made by Ilse)

Other low-hanging fruit: structure notebook entries

Box 1: Data storage

Box 2: Aim

Box 3: Samples

Last box: Conclusions

20181217_FMH-114_NUTM1 stain cytopins

Project:	b_Chr21	Status:	Completed 
Study:	NUTM1	ExperimentID:	001000000000170297
Created By:	Femke Hormann	Created:	2018-12-21 18:57
		Due date:	not set 

[Toggle collapse all sections](#)

 Data storage   

2018-12-21

L:\pmc_denboer\fhormann\elabjournal\Chr21\NUTM1\20181217_FMH-114_NUTM1 stain cytopins

created: 2018-12-21 Femke Hormann

last updated: 2018-12-21 Femke Hormann

 Introduction & objectives   

2018-12-21

Introduction

We have received feedback on our NUTM1 paper from the reviewers. A discussion point was that we didn't show protein expression of NUTM1 and that we didn't show the cellular localization. Therefore we are staining cytopins for NUTM1 to show the expression and the cellular localization.

Objectives

Staining all NUTM1 fusion samples with the NUTM1 antibody to see (nuclear) staining.

created: 2018-12-21 Femke Hormann

last updated: 2018-12-21 Femke Hormann

 Samples   

2018-12-21

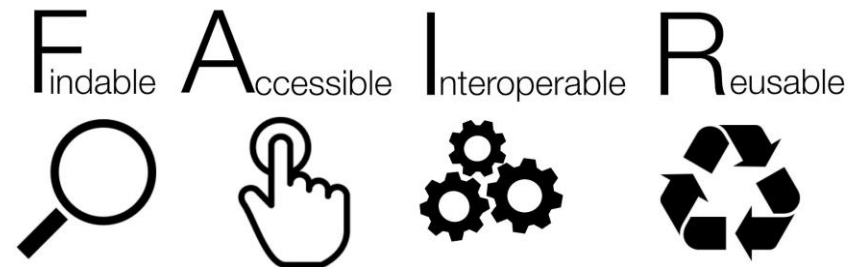
The secret: continuous attention and stimulation

- Importance stressed by PI: *group culture*
- Bottom-up practical approach: *group effort*
- Group data steward and quality team: *group structure*
- Monthly work discussions from lab notebook: *group pressure*
- Introduction to new employees: *group building*
- **Experiencing in practice that it works!**



And finally: how FAIR are we?

- **Findable:** transparent file structure linked to lab notebook
- **Accessible:** for group members; historic data via group data steward
- **Interoperable:** data format indicated per file type
- **Reusable:** for group members, for collaborators, public repositories



ZonMW Open Science Impuls 2019 awarded to Monique den Boer



“Data stewardship is a team effort with high gains in data transparency and reproducibility.”

“Structuring experimental data takes a bit of time to set up, but then the time investment progressively decreases while the scientific benefit impressively increases!”



**Oncode
Institute**



**princess
máxima
center**
pediatric oncology

**Thank you
for listening!**



j.m.boer-20@prinsesmaximacentrum.nl